

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2005, 23:42:26 ; Search time 81.1765 Seconds
(without alignments)
61.798 Million cell updates/sec

Title: US-10-766-752B-9

Perfect score: 57

Sequence: 1 DXCXGDSGGPXV 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	53	93.0	248	20	US-11-097-143-22677
2	53	93.0	251	11	US-09-789-210-47
3	53	93.0	252	14	US-10-148-671-21
4	53	93.0	252	20	US-11-097-143-17229
5	53	93.0	282	15	US-10-051-874-97
6	53	93.0	296	14	US-10-148-671-17
7	53	93.0	322	15	US-10-051-874-96
8	53	93.0	362	20	US-11-097-143-12843
9	53	93.0	477	20	US-11-097-143-24180
10	53	93.0	560	20	US-11-097-143-8433
11	53	93.0	611	20	US-11-097-143-24174

Sequence 6, Appli
Sequence 6, Appli
Sequence 2, Appli
Sequence 93, Appli
Sequence 148, Appli
Sequence 38814, A
Sequence 4, Appli
Sequence 8, Appli
Sequence 4, Appli
Sequence 8, Appli
Sequence 6, Appli
Sequence 2, Appli
Sequence 6, Appli
Sequence 14232, A
Sequence 2, Appli
Sequence 33, Appli
Sequence 34, Appli
Sequence 18, Appli
Sequence 19, Appli
Sequence 53, Appli
Sequence 2, Appli
Sequence 18, Appli
Sequence 19, Appli
Sequence 33, Appli
Sequence 34, Appli
Sequence 33, Appli
Sequence 34, Appli
Sequence 2, Appli
Sequence 10, Appli
Sequence 14, Appli
Sequence 16, Appli
Sequence 18, Appli

679 9 US-09-874-198-6
679 9 US-09-874-238-6
679 20 US-11-032-149-6
699 15 US-10-388-322-2
699 18 US-10-820-155-93
704 18 US-10-989-891-148
774 20 US-11-097-143-38814
1019 14 US-10-183-992-4
1019 14 US-10-183-992-8
1019 16 US-10-638-125-4
1019 16 US-10-480-254-4
1019 16 US-10-480-254-8
1083 14 US-10-183-992-6
1083 16 US-10-638-125-2
1083 16 US-10-480-254-6
1190 20 US-11-097-143-14232
1190 20 US-09-904-090-2
12 9 US-09-804-156-33
12 9 US-09-804-156-34
12 9 US-09-946-633-18
12 9 US-09-946-633-19
12 10 US-09-997-003-53
12 13 US-10-050-611-2
12 13 US-10-125-459-18
12 13 US-10-125-459-19
12 13 US-10-067-761-33
12 13 US-10-067-761-34
12 14 US-10-319-519-33
12 14 US-10-319-519-34
12 18 US-10-957-427-2
12 18 US-10-843-299-10
13 17 US-10-843-299-14
13 17 US-10-843-299-16
13 17 US-10-843-299-18

ALIGNMENTS

RESULT 1

US-11-097-143-22677
; Sequence 22677, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22677
; LENGTH: 248
; TYPE: PRT

! ORGANISM: DROSOPHILA
US-11-097-143-22677

Query Match 93.0%; Score 53; DB 20; Length 248;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
DB 192 DACSGDSGGPLV 203

RESULT 2

US-09-789-210-47
; Sequence 47, Application US/09789210
; Publication No. US20040241846A1
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; COLPITTS, TRACEY L.
; FRIEDMAN, PAULA N.
; GRANADOS, EDWARD N.
; KLASS, MICHAEL R.
; RUSSELL, JOHN C.
; STEWART, KENT D.
; STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/789,210
; FILING DATE: 20-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/944,483
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-3623
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 251 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:

Query Match 93.0%; Score 53; DB 11; Length 251;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
DB 192 DACSGDSGGPMV 203

RESULT 3

US-10-148-671-21
; Sequence 21, Application US/10148671
; Publication No. US20030186419A1
; GENERAL INFORMATION:
; APPLICANT: Jensenius, Jens Christian
; APPLICANT: Thiel, Steffen
; TITLE OF INVENTION: MASP-3, A complement-fixing enzyme, and uses for it
; FILE REFERENCE: 10/148,671
; CURRENT APPLICATION NUMBER: US/10/148,671
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: PCT/DK00/00659
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 21
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens (fig. 10, huMASP-1)
US-10-148-671-21

Query Match 93.0%; Score 53; DB 14; Length 252;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
DB 193 DACAGDSGGPMV 204

RESULT 4

US-11-097-143-17229
; Sequence 17229, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17229
; LENGTH: 252
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-17229

Query Match 93.0%; Score 53; DB 20; Length 252;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | |
Db 200 DACTGDSGGPLV 211

RESULT 5

US-10-051-874-97
; Sequence 97, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glennda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Herrman, John L
; APPLICANT: Payman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS. POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; PRIOR FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265,530
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/261,376
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 97
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-051-874-97

Query Match 93.0%; Score 53; DB 15; Length 282;
Best Local Similarity 75.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | |
Db 200 DTCGDSGGPLV 211

RESULT 6

US-10-148-671-17
; Sequence 17, Application US/10148671
; Publication No. US20030186419A1
; GENERAL INFORMATION:
; APPLICANT: Jensenius, Jens Christian
; APPLICANT: Thiel, Steffen
; TITLE OF INVENTION: MASP-3, A complement-fixing enzyme, and uses for it
; FILE REFERENCE: 10/148,671
; CURRENT APPLICATION NUMBER: US/10/148,671
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: PCT/DK00/00659
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Homo sapiens (fig. 6, MASP-1)
US-10-148-671-17

Query Match 93.0%; Score 53; DB 14; Length 296;
Best Local Similarity 75.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | |
Db 237 DACAGDSGGPMV 248

RESULT 7

US-10-051-874-96
; Sequence 96, Application US/10051874
; Publication No. US20040005557A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Alsobrook II, John P
; APPLICANT: Colman, Steven D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Edinger, Shlomit R
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Pena, Carol EA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Millet, Isabelle
; APPLICANT: Miller, Charles E
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glennda
; APPLICANT: Baumgartner, Jason C

```

; APPLICANT: Herrman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Burgess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; PRIOR FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265,530
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/261,376
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 269
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 96
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-051-874-96

Query Match          93.0%; Score 53; DB 15; Length 322;
Best Local Similarity 75.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 DXCXGDSGGPXV 12
DB      240 DTCXGDSGGPLV 251

RESULT 8
US-11-097-143-12843
; Sequence 12843, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24180
; LENGTH: 477
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-24180

Query Match          93.0%; Score 53; DB 20; Length 477;
Best Local Similarity 75.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 DXCXGDSGGPXV 12
DB      201 DSCXGDSGGPLV 212

```

```

; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12843
; LENGTH: 362
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-12843

```

```

Query Match          93.0%; Score 53; DB 20; Length 362;
Best Local Similarity 75.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 DXCXGDSGGPXV 12
DB      239 DACAGDSGGPLV 250

```

```

RESULT 9
US-11-097-143-24180
; Sequence 24180, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24180
; LENGTH: 477
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-24180

```

```

Query Match          93.0%; Score 53; DB 20; Length 477;
Best Local Similarity 75.0%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 DXCXGDSGGPXV 12
DB      201 DSCXGDSGGPLV 212

```

```
RESULT 10
US-11-097-143-8433
; Sequence 8433, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728.
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8433
; LENGTH: 580
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-8433

Query Match          93.0%; Score 53; DB 20; Length 580;
Best Local Similarity 75.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
   | | | | | | | |
Db 528 DTCXGDSGGPLV 539

RESULT 11
US-11-097-143-24174
; Sequence 24174, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8433
; LENGTH: 580
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-8433

Query Match          93.0%; Score 53; DB 20; Length 580;
Best Local Similarity 75.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
   | | | | | | | |
Db 528 DTCXGDSGGPLV 539
```

```
RESULT 12
US-09-874-198-6
; Sequence 6, Application US/09874198
; Patent No. US20020082208A1
; GENERAL INFORMATION:
; APPLICANT: Jensenius, Jens Chr.
; APPLICANT: Thiel, Steffen
; TITLE OF INVENTION: MASP-2 COMPLEMENT-FIXING ENZYME, AND
; TITLE OF INVENTION: USES FOR IT
; FILE REFERENCE: 09011-002002
; CURRENT APPLICATION NUMBER: US/09/874,198
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/054,218
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 60/042,678
; PRIOR FILING DATE: 1997-04-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-874-198-6

Query Match          93.0%; Score 53; DB 9; Length 679;
Best Local Similarity 75.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
   | | | | | | | |
Db 620 DACSGDSGGPMV 631

RESULT 13
US-09-874-238-6
; Sequence 6, Application US/09874238
; Patent No. US20020082209A1
; GENERAL INFORMATION:
; APPLICANT: Jensenius, Jens Chr.
; APPLICANT: Thiel, Steffen
; TITLE OF INVENTION: MASP-2 COMPLEMENT-FIXING ENZYME, AND
; TITLE OF INVENTION: USES FOR IT
; FILE REFERENCE: 09011-002003
; CURRENT APPLICATION NUMBER: US/09/874,238
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/054,218
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 60/042,678
; PRIOR FILING DATE: 1997-04-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 679
```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-874-239-6

Query Match 93.0%; Score 53; DB 9; Length 679;
Best Local Similarity 75.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | |
DB 620 DACSGDSGGPMV 631

RESULT 14

US-11-032-149-6
; Sequence 6, Application US/11032149
; Publication No. US20050158297A1
; GENERAL INFORMATION:
; APPLICANT: Jensenius, Jens Chr.
; TITLE OF INVENTION: MASP-2 COMPLEMENT-FIXING ENZYME, AND
; FILE OF INVENTION: USES FOR IT
; FILE REFERENCE: 09011-002002
; CURRENT APPLICATION NUMBER: US/11/032,149
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: US/09/874,198
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/054,218
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 60/042,678
; PRIOR FILING DATE: 1997-04-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-032-149-6

Query Match 93.0%; Score 53; DB 20; Length 679;
Best Local Similarity 75.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | |
DB 620 DACSGDSGGPMV 631

RESULT 15

US-10-388-322-2
; Sequence 2, Application US/10388322
; Publication No. US20040006009A1
; GENERAL INFORMATION:
; APPLICANT: NatImmune
; TITLE OF INVENTION: Pharmaceutical compositions comprising mannose binding lectin
; FILE REFERENCE: P 625 DK00
; CURRENT APPLICATION NUMBER: US/10/388,322
; CURRENT FILING DATE: 2003-03-14
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-388-322-2

Query Match 93.0%; Score 53; DB 15; Length 699;
Best Local Similarity 75.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | |
DB 640 DACSGDSGGPMV 651

Search completed: November 2, 2005, 00:12:57
Job time : 82.1765 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2005, 23:33:46 ; Search time 25.1765 Seconds
(without alignments)
35.580 Million cell updates/sec

Title: US-10-766-752B-9

Perfect score: 57

Sequence: 1 DXCXGDSGGPXV 12

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	93.0	251	3	US-08-944-483-47
2	53	93.0	699	4	US-09-949-016-6138
3	53	93.0	717	4	US-09-949-016-11182
4	53	93.0	1019	1	US-08-296-014A-4
5	53	93.0	1019	2	US-08-596-405-4
6	53	93.0	1019	2	US-08-877-620-4
7	53	93.0	1019	2	US-08-877-368-4
8	53	93.0	1019	4	US-09-626-795-4
9	53	93.0	1083	1	US-08-296-014A-2
10	53	93.0	1083	2	US-08-596-405-2
11	53	93.0	1083	2	US-08-877-620-2
12	53	93.0	1083	4	US-09-287-368-2
13	53	93.0	1083	4	US-09-626-795-2
14	52	91.2	12	4	US-08-538-504-2
15	52	91.2	12	4	US-09-631-137C-2
16	52	91.2	12	6	5352664-3
17	52	91.2	12	6	5352664-3
18	52	91.2	13	4	US-09-108-006C-34
19	52	91.2	14	4	US-09-631-137C-7
20	52	91.2	14	4	US-09-909-348A-4
21	52	91.2	23	3	US-08-981-088-2
22	52	91.2	23	3	US-08-330-594-8
23	52	91.2	23	4	US-09-644-038-8
24	52	91.2	23	4	US-08-538-504-3
25	52	91.2	23	4	US-09-631-137C-3
26	52	91.2	23	4	US-09-879-792-22
27	52	91.2	23	4	US-09-909-348A-5

28	52	91.2	23	4	US-09-909-348A-6
29	52	91.2	138	6	5200340-4
30	52	91.2	138	6	5200340-4
31	52	91.2	149	3	US-09-518-046-20
32	52	91.2	151	3	US-09-518-046-21
33	52	91.2	151	4	US-09-270-767-33178
34	52	91.2	151	4	US-09-270-767-33178
35	52	91.2	157	3	US-09-518-046-23
36	52	91.2	158	3	US-09-518-046-22
37	52	91.2	159	3	US-09-518-046-24
38	52	91.2	162	4	US-09-244-111-6
39	52	91.2	164	3	US-09-518-046-25
40	52	91.2	182	4	US-09-328-925-12
41	52	91.2	197	1	US-08-456-840-48
42	52	91.2	197	1	US-08-266-407A-48
43	52	91.2	197	2	US-08-892-544-48
44	52	91.2	207	4	US-09-244-111-4
45	52	91.2	207	4	US-10-000-489-54

ALIGNMENTS

RESULT 1

US-08-944-483-47

; Sequence 47, Application US/08944483

; Patent No. 6232456

; GENERAL INFORMATION:

; APPLICANT: COHEN, MAURICE

; APPLICANT: COLPITTS, TRACEY L.

; APPLICANT: FRIEDMAN, PAULA N.

; APPLICANT: GRANADOS, EDWARD N.

; APPLICANT: KLASS, MICHAEL R.

; APPLICANT: RUSSELL, JOHN C.

; APPLICANT: STEWART, KENT D.

; APPLICANT: STROUPE, STEVEN D.

; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

; NUMBER OF SEQUENCES: 76

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Abbott Laboratories

; STREET: 100 Abbott Park Road

; CITY: Abbott Park

; STATE: IL

; COUNTRY: USA

; ZIP: 60064-3500

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/944,483

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Becker, Cheryl L.

; REGISTRATION NUMBER: 35,441

; REFERENCE/DOCKET NUMBER: 6183.US.01

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 847/935-1729

; TELEFAX: 847/938-2623

; TELEX:

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 251 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

```
; MOLECULE TYPE: No. 6232456e
US-08-944-483-47

Query Match      93.0%; Score 53; DB 3; Length 251;
Best Local Similarity 75.0%; Pred. No. 0.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
Db 192 DACSGDSGGPMV 203

RESULT 2
US-09-949-016-6138
; Sequence 6138, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6138
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6138

Query Match      93.0%; Score 53; DB 4; Length 699;
Best Local Similarity 75.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
Db 640 DACSGDSGGPMV 651

RESULT 3
US-09-949-016-11182
; Sequence 11182, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11182
; LENGTH: 717
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11182

Query Match      93.0%; Score 53; DB 4; Length 717;
Best Local Similarity 75.0%; Pred. No. 1;

; MOLECULE TYPE: No. 6232456e
US-08-944-483-47

Query Match      93.0%; Score 53; DB 1; Length 1019;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
Db 960 DACSGDSGGPLV 971

RESULT 5
US-08-596-405-4
; Sequence 4, Application US/08596405
; Patent No. 5858706
; GENERAL INFORMATION:
; APPLICANT: Ho, Bow
; TITLE OF INVENTION: The Cloned Factor C cDNA of the
; ADDRESSSEE: Singapore Horseshoe Crab, Carcinoscopus
; TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,014A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1781-105P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1019 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-296-014A-4

Query Match      93.0%; Score 53; DB 1; Length 1019;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
Db 960 DACSGDSGGPLV 971

RESULT 5
US-08-596-405-4
; Sequence 4, Application US/08596405
; Patent No. 5858706
; GENERAL INFORMATION:
; APPLICANT: Ho, Bow
; TITLE OF INVENTION: The Cloned Factor C cDNA of the
; ADDRESSSEE: Singapore Horseshoe Crab, Carcinoscopus
; TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
```


CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/596,405
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1781-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1019 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-596-405-4

Query Match 93.0%; Score 53; DB 2; Length 1019;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | | | |
Db 960 DACSGDSGGPLV 971

RESULT 6
US-08-877-620-4
Sequence 4, Application US/08877620
Patent No. 5985590
GENERAL INFORMATION:
APPLICANT: Ding, Jeak Ling
APPLICANT: HO, Bow
TITLE OF INVENTION: The Cloned Factor C cDNA of the
TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscorpis
TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/877,620
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/596,405
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1781-105P
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1019 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-877-620-4

Query Match 93.0%; Score 53; DB 2; Length 1019;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | | | |
Db 960 DACSGDSGGPLV 971

RESULT 7
US-09-287-368-4
Sequence 4, Application US/09287368A
Patent No. 6645724
GENERAL INFORMATION:
APPLICANT: DING, Jeak Ling
APPLICANT: HO, Bow
TITLE OF INVENTION: Assays for Endotoxin and Methods for Removal of Endotoxin
TITLE OF INVENTION: from a Sample Using Recombinant Factor C
FILE REFERENCE: 1781-0165P
CURRENT APPLICATION NUMBER: US/09/287,368A
CURRENT FILING DATE: 1999-04-07
EARLIER APPLICATION NUMBER: 09/201,786
EARLIER FILING DATE: 1998-12-01
EARLIER APPLICATION NUMBER: 09/081,767
EARLIER FILING DATE: 1998-05-21
EARLIER APPLICATION NUMBER: 60/058,816
EARLIER FILING DATE: 1997-09-19
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 4
LENGTH: 1019
TYPE: PRT
ORGANISM: Carcinoscorpis rotundicauda
FEATURE:
OTHER INFORMATION: any n or Xaa = Unknown
US-09-287-368-4

Query Match 93.0%; Score 53; DB 4; Length 1019;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | | | |
Db 960 DACSGDSGGPLV 971

RESULT 8
US-09-626-795-4
Sequence 4, Application US/09626795
Patent No. 6719973
GENERAL INFORMATION:
APPLICANT: DING, Jeak Ling
APPLICANT: HO, Bow
TITLE OF INVENTION: Use of Recombinant Factor C to Induce Bacteriostasis
FILE REFERENCE: 4810-61737
CURRENT APPLICATION NUMBER: US/09/626,795
CURRENT FILING DATE: 2000-07-26
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 4
LENGTH: 1019
TYPE: PRT

```
; ORGANISM: Carcinoscopus rotundicauda
US-09-626-795-4

Query Match          93.0%; Score 53; DB 4; Length 1019;
Best Local Similarity 75.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
Db 960 DACSGDSGGPLV 971

RESULT 9
US-08-296-014A-2
; Sequence 2, Application US/08296014A
; Patent No. 5716834
; GENERAL INFORMATION:
; APPLICANT: Ding, Jeak Ling
; APPLICANT: Ho, Bow
; TITLE OF INVENTION: The Cloned Factor C cDNA of the
; TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscopus
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,014A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy, Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1781-105P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1083 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-296-014A-2

Query Match          93.0%; Score 53; DB 1; Length 1083;
Best Local Similarity 75.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
Db 1024 DACSGDSGGPLV 1035

RESULT 10
US-08-596-405-2
; Sequence 2, Application US/08596405
; Patent No. 5858706
; GENERAL INFORMATION:
; APPLICANT: Ding, Jeak Ling
; APPLICANT: Ho, Bow
; TITLE OF INVENTION: The Cloned Factor C cDNA of the
; TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscopus
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/877,620
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 08/596,405
; FILING DATE:
```

ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1781-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1083 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-877-620-2

Query Match 93.0%; Score 53; DB 2; Length 1083;
Best Local Similarity 75.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGP XV 12
| | | | | | | | | |
Db 1024 DACSGDSGGPLV 1035

RESULT 12
US-09-287-368-2
Sequence 2, Application US/09287368A
Patent No. 6645724
GENERAL INFORMATION:
APPLICANT: DING, Jeak Ling
APPLICANT: HO, Bow
TITLE OF INVENTION: Assays for Endotoxin and Methods for Removal of Endotoxin
TITLE OF INVENTION: from a Sample Using Recombinant Factor C
FILE REFERENCE: 1781-0165P
CURRENT APPLICATION NUMBER: US/09/287,368A
CURRENT FILING DATE: 1999-04-07
EARLIER APPLICATION NUMBER: 09/201,786
EARLIER FILING DATE: 1998-12-01
EARLIER APPLICATION NUMBER: 09/081,767
EARLIER FILING DATE: 1998-05-21
EARLIER APPLICATION NUMBER: 60/058,816
EARLIER FILING DATE: 1997-09-19
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
TYPE: PRT
ORGANISM: Carcinoscorpius rotundicauda
FEATURE:
OTHER INFORMATION: any n or Xaa = Unknown
US-09-287-368-2

Query Match 93.0%; Score 53; DB 4; Length 1083;
Best Local Similarity 75.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGP XV 12
| | | | | | | | | |
Db 1024 DACSGDSGGPLV 1035

RESULT 13
US-09-626-795-2
Sequence 2, Application US/09626795
Patent No. 6719973
GENERAL INFORMATION:
APPLICANT: DING, Jeak Ling
APPLICANT: HO, Bow
TITLE OF INVENTION: Use of Recombinant Factor C to Induce Bacteriostasis
FILE REFERENCE: 4810-61737
CURRENT APPLICATION NUMBER: US/09/626,795
CURRENT FILING DATE: 2000-07-26

NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1083
TYPE: PRT
ORGANISM: Carcinoscorpius rotundicauda
US-09-626-795-2

Query Match 93.0%; Score 53; DB 4; Length 1083;
Best Local Similarity 75.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGP XV 12
| | | | | | | | | |
Db 1024 DACSGDSGGPLV 1035

RESULT 14
US-08-538-504-2
Sequence 2, Application US/08538504
Patent No. 6627731
GENERAL INFORMATION:
APPLICANT: Carney, Darrell H.
APPLICANT: Glenn, Kevin C.
TITLE OF INVENTION: THROMBIN DERIVED POLYPEPTIDES:
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR USE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/538,504
FILING DATE: Concurrently Herewith
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSG:178/PAR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-538-504-2

Query Match 91.2%; Score 52; DB 4; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.036;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGP XV 12
| | | | | | | | | |
Db 1 DACEGDSGGP XV 12

RESULT 15
US-09-631-137C-2
Sequence 2, Application US/09631137C
Patent No. 6630572
GENERAL INFORMATION:

```

; APPLICANT: Carney, Darrell H.
; APPLICANT: Glenn, Kevin C.
; TITLE OF INVENTION: Thrombin Derived Polypeptides:
; TITLE OF INVENTION: Compositions and Methods for Use
; FILE REFERENCE: 3033.1001-004
; CURRENT APPLICATION NUMBER: US/09/631,137C
; CURRENT FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 08/538,504
; PRIOR FILING DATE: 1995-09-29
; PRIOR APPLICATION NUMBER: US 08/007,173
; PRIOR FILING DATE: 1993-01-21
; PRIOR APPLICATION NUMBER: US 06/925,201
; PRIOR FILING DATE: 1986-10-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fragment of human prothrombin
US-09-631-137C-2

```

```

Query Match          91.2%; Score 52; DB 4; Length 12;
Best Local Similarity 75.0%; Pred. No. 0.036;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 DXCXGDSGGPFV 12
   | | | | | | |
Db 1 DACEGDSGGPFV 12

```

```

Search completed: November 1, 2005, 23:45:00
Job time : 25.1765 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2005, 23:21:14 ; Search time 47.2941 Seconds
(without alignments)
129.931 Million cell updates/sec

Title: US-10-766-752B-9
Perfect score: 57
Sequence: 1 DXCXGDSGGPXV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03: *
1: uniprot_prot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	94.7	258	2 O97399	O97399 phaeton coc
2	54	94.7	263	2 Q7PUB9	Q7PUB9 anopheles g
3	53	93.0	248	2 O16126	O16126 boltenia vi
4	53	93.0	248	2 Q8IRE2	Q8IRE2 drosophila
5	53	93.0	248	2 Q9VQ98	Q9VQ98 drosophila
6	53	93.0	252	1 TRVI DROME	P52905 drosophila
7	53	93.0	257	2 Q9VZT0	Q9VZT0 drosophila
8	53	93.0	259	2 Q8IRE0	Q8IRE0 drosophila
9	53	93.0	260	2 Q6VPUE	Q6VPUE sarcophag s
10	53	93.0	270	2 Q641D1	Q641D1 anthonomus
11	53	93.0	277	2 Q8MLC4	Q8MLC4 drosophila
12	53	93.0	282	2 Q9D4I3	Q9D4I3 mus musculu
13	53	93.0	296	2 Q8T9UE	Q8T9UE aedes aegypt
14	53	93.0	299	2 Q9VS87	Q9VS87 drosophila
15	53	93.0	322	2 Q920S2	Q920S2 mus musculu
16	53	93.0	345	2 Q96088	Q96088 haemaphysal
17	53	93.0	362	2 Q9W453	Q9W453 drosophila
18	53	93.0	419	2 Q688K5	Q688K5 nanochlori
19	53	93.0	483	2 Q8T8X4	Q8T8X4 drosophila
20	53	93.0	483	2 Q9VK10	Q9VK10 drosophila
21	53	93.0	536	2 Q7PX72	Q7PX72 anopheles g
22	53	93.0	698	2 Q6GPF9	Q6GPF9 xenopus lae
23	53	93.0	698	2 Q9PU71	Q9PU71 xenopus lae
24	53	93.0	699	1 CRAR HUMAN	P48740 h complemen
25	53	93.0	701	2 Q9UJ59	Q9UJ59 rattus norv
26	53	93.0	703	2 Q8CHN8	Q8CHN8 rattus norv
27	53	93.0	704	1 CRAR MOUSE	P98064 mus musculu
28	53	93.0	1019	1 LFC CARRO	Q26422 carcinoscor
29	53	93.0	1019	1 LFC TACTR	P28175 tachyleus
30	53	93.0	1019	2 Q8T9S1	Q8T9S1 tachyleus
31	53	93.0	1083	2 Q26423	Q26423 carcinoscor

32	52	91.2	54	2 Q6JDI3	Q6JDI3 canis famil
33	52	91.2	66	2 Q61752	Q61752 mus musculu
34	52	91.2	73	2 Q9TV90	Q9TV90 equus cabal
35	52	91.2	83	2 Q6JVS2	Q6JVS2 sus scrofa
36	52	91.2	85	2 Q8MVL1	Q8MVL1 boltenia vi
37	52	91.2	87	2 Q9CQ78	Q9CQ78 m mus muscu
38	52	91.2	117	2 Q9PUF3	Q9PUF3 bothrops ja
39	52	91.2	119	2 Q9NR68	Q9NR68 homo sapien
40	52	91.2	124	2 Q8C6G5	Q8C6G5 mus musculu
41	52	91.2	125	2 Q86VI8	Q86VI8 homo sapien
42	52	91.2	125	2 Q804G0	Q804G0 spherooides
43	52	91.2	136	2 Q6TG90	Q6TG90 drosophila
44	52	91.2	141	2 Q6GKZ6	Q6GKZ6 drosophila
45	52	91.2	142	2 Q8HYM3	Q8HYM3 felis silve

ALIGNMENTS

RESULT 1
O97399 PRELIMINARY; PRT; 258 AA.
ID O97399;
AC O97399;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Trypsin precursor.
OS Phaeton cochleariae (Mustard beetle).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytomphaga; Chrysomeloidea; Chrysomelidae; Chrysomelinae;
OC Chrysomelini; Phaeton.
NCBI_TaxID=80249;
RN [1] RN
RX
RP SEQUENCE FROM N.A.
RC
RC Tissue=Gut;
CA Girard C., Jouanin L.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; Y17905; CAA76929.1; -;
DR HSP; P00760; LE2X
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001354; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPG; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN 1.
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 16 potential.
FT CHAIN 30 258 trypsin.
SQ SEQUENCE 258 AA; 28070 MW; BDBDFAFECB86866C CRC64;

Query Match 94.7%; Score 54; DB 2; Length 258;
Best Local Similarity 75.0%; Pred. No. 0.12; Mismatches 0; Gaps 0;
Matches 9; Conservative 0; Indels 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | | | |
Db 207 DSCXGDSGGPAV 218

RESULT 2
Q7PUB9 PRELIMINARY; PRT; 263 AA.
ID Q7PUB9
AC Q7PUB9;
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE ENSANGP00000013856 (Fragment).
GN Name=ENSANGG00000011367.
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008987; EAA01780.2; -.
DR HSP; P00736; 1GPZ.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1_Ser_Cys.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
SQ SEQUENCE 263 AA; 28229 MW; 1054300AEBE4BCE1 CRC64;
Query Match 94.7%; Score 54; DB 2; Length 263;
Best Local Similarity 75.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 DXCXGDSGGPXV 12
DB 210 DACAGDSGGPAV 221
RESULT 3
O16126 PRELIMINARY; PRT; 248 AA.
ID O16126
AC O16126
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Trypsinogen 1 precursor.
GN Name=TRYPI;
OS Boltenia villosa.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyruidae; Boltenia.
OX NCBI_TaxID=63515;
RN [1]
RP SEQUENCE FROM N.A.
RA Roach J.C.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AF011897; AAB69653.1; -.
DR HSP; P00760; 1EZK.
DR MEROPS; S01.130; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.

DR PROSITE; PS02440; TRYPSIN_DOM; 1.
KW Hydrolase; Protease; Serine protease; Signal.
FT SIGNAL 1 12 Potential.
FT CHAIN 13 248 tryDsein 1.
SQ SEQUENCE 248 AA; 25872 MW; AC06B8998413305 CRC64;
Query Match 93.0%; Score 53; DB 2; Length 248;
Best Local Similarity 75.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 DXCXGDSGGPXV 12
DB 198 DSCQDGGGPAV 209
RESULT 4
Q8IRE2 PRELIMINARY; PRT; 248 AA.
ID Q8IRE2
AC Q8IRE2;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE CG32271-PA.
GN ORFNames=CG32271;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherzer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hootin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RN Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RG FLYBASE;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RG FLYBASE;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AF003583; AAF51279.1; -.
DR HSSP; P00750; 1RTF.
DR FLYBASE; FBgn0042186; CG17239.
DR GO; GO:0004295; F:trypsin activity; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 248 AA, 26744 MW, 0780FD663F005807 CRC64;

Query Match 93.0%; Score 53; DB 2; Length 248;
Best Local Similarity 75.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPVX 12
| | | | | | | |
DB 192 DACSGDSGGPLV 203

RESULT 6
TRYI_DROME
ID TRYI_DROME STANDARD; PRT; 252 AA.
AC P52905; Q9V5Y6;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Trypsin iota precursor (EC 3.4.21.4).
GN Name=iota-Try; Synonyms=TRY-IOTA; ORFNames=CG7754;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RC Gao L., Wang S., Hickey D.A.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.A., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.R., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RN Science 287:2185-2195(2000).
RN [3]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Shu S.E.;
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: Belongs to the peptidase S1 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U04853; AAA84377.1; -.
CC EMBL; AF003826; AAF58655.1; -.
CC HSSP; P00760; 3BTH.
CC MEROPS; S01.115; -.
CC FLYBASE; FBgn0015001; iota-Try.
CC GO; GO:0004295; F:trypsin activity; NAS.
CC GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
CC InterPro; IPR009003; Pept_Ser_Cys.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00089; Trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; TRYP_SPC; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Multiagent family; Serine protease; Signal; Zymogen.
KW

FT SIGNAL 1 19 Potential.
 FT CHAIN 20 27 Activation peptide.
 FT CHAIN 28 252 Trypsin Iota.
 FT ACT_SITE 68 68 Charge relay system (By similarity).
 FT ACT_SITE 113 113 Charge relay system (By similarity).
 FT ACT_SITE 206 206 Charge relay system (By similarity).
 FT DISULFID 53 69 By similarity.
 FT DISULFID 175 193 By similarity.
 FT DISULFID 202 226 By similarity.
 FT SITE 200 200 Required for specificity (By similarity).
 SQ SEQUENCE 252 AA; 26612 MW; 044DEDA4700910C2 CRC64;
 Query Match 93.0%; Score 53; DB 1; Length 252;
 Best Local Similarity 75.0%; Pred. No. 0.17;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 DXCXGDSGGPXV 12
 Db 200 DACTGDSGGPLV 211
 ID QSVZT0 PRELIMINARY; PRT; 257 AA.
 AC QSVZT0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE CG33159-PA.
 GN ORFNames=CG33159;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Jalaali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasmann D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."; Science 287:2185-2195(2000).
 RL [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22426085; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila melanogaster euchromatic genome sequence."; Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RL [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin: a genomic perspective."; Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RL [4]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Mitra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a systematic review."; Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RL [5]
 RN SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE FROM N.A.
 RP FlyBase;
 RG Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 RL -1- SIMILARITY: Belongs to peptidase family S1.
 CC EMBL; AE003477; AAF4737.2; -.
 DR HSP; P00761; IEPT.
 DR FlyBase; Pfam0052271; CG32271.
 DR GO; GO:0004295; F:trypsin activity; NAS.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
 DR InterPro; IPR01254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYP_SPC; 1.
 DR PROSITE; PS00240; TRYP_SIN_DOM; 1.
 DR PROSITE; PS00134; TRYP_SIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYP_SIN_SER; 1.
 KW Hydrolase; Serine protease.
 SQ SEQUENCE 257 AA; 27795 MW; BFFB09E0ED122691 CRC64;
 Query Match 93.0%; Score 53; DB 2; Length 257;
 Best Local Similarity 75.0%; Pred. No. 0.18;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 DXCXGDSGGPXV 12
 Db 201 DSCSGDSGGPLV 212
 RESULT 8

Q8IRE0 PRELIMINARY; PRT; 259 AA.
 AC Q8IRE0;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE CG3161-PA.
 GN ORFNames=CG3161;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier A., Fleischmann W.,
 RA Fostler C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skuiski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.O.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun. Release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence."
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:

RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 [4]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 [5]
 RN SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 [6]
 RN SEQUENCE FROM N.A.
 RP FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to peptidase family S1.
 DR EMBL; AE003477; AAN11545.1; -
 DR HSP; P00750; 1RTP
 DR FlyBase; FBgn0052270; CG32270.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001254; Peptidase_S1.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM0020; Tryp_Spc; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 259 AA; 28699 MW; 070BDAB8259277E CRC64;
 Query Match 93.0%; Score 53; DB 2; Length 259;
 Best Local Similarity 75.0%; Pred. No. 0.18; Mismatches 0; Gaps 0;
 Matches 9; Conservative 0; Indels 3; Indels 0; Gaps 0;
 QY 1 DXCGDGGGPPV 12
 DB 203 DACAGDGGPPV 214
 RESULT 9
 Q6VPUG PRELIMINARY; PRT; 260 AA.
 AC Q6VPUG;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Sar s 3 allergen Yv7016G03.
 OS Sarcopes scabiei type hominis.
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
 OC Acariformes; Sarcopiformes; Astigmata; Psoroptidia; Sarcopitoidea;
 OC Sarcopitidae; Sarcopes.
 OC NCBI_TaxID=197185;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14675192;
 RA Holt D.C., Fischer K., Allen G.E., Wilson D., Wilson P., Slade R.,
 RA Currie B.J., Walton S.F., Kemp D.J.;
 RT "Mechanisms for a novel immune evasion strategy in the scabies mite
 RT sarcopes scabiei: a multigene family of inactivated serine
 RT proteases.";

RL J. Invest. Dermatol. 121:1419-1424(2003).
CC -I- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AY333071; ARI14081.1; -
DR HSP; P00761; IAKS.
DR GO; GO:0004263; F:Chymotrypsin activity; IEA.
DR GO; GO:0008233; F:Peptidase activity; IEA.
DR GO; GO:0004295; F:Trypsin activity; IEA.
DR GO; GO:0006508; P:Proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1-
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 260 AA; 29085 MW; DD9B5B8C43D6D467 CRC64;
Query Match 93.0%; Score 53; DB 2; Length 260;
Best Local Similarity 75.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 DXCKGDSGGPXV 12
Db 208 DSCRGDSGGPVP 219
RESULT 10
ID Q641D1 PRELIMINARY; PRT; 270 AA.
AC Q641D1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Trypsin-like serine proteinase.
GN Name-ser21;
OS Anthonomus grandis (Boll weevil).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Phytophaga; Curculionidae; Curculioninae; Anthonomini; Anthonomus.
OX NCBI_TaxID=7044;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15350610;
RA Oliveira-Neto O.B., Batista J.A.N., Rigden D.J., Fragoso R.R.,
RA Silva R.O., Gomes E.A., Franco O.L., Dias S.C., Cordeiro C.M.T.,
RA Monnerat R.G., Grossi-De-Sa M.F.;
RT "A diverse family of serine proteinase genes expressed in cotton boll
weevil (Anthonomus grandis): implications for the design of pest-
resistant transgenic cotton plants.";
RL Insect Biochem. Mol. Biol. 34:903-918(2004).
DR EMBL; AY536264; AAT09850.1; -
SQ SEQUENCE 270 AA; 28267 MW; 3EAB7B85AF4E4A3A CRC64;
Query Match 93.0%; Score 53; DB 2; Length 270;
Best Local Similarity 75.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 DXCKGDSGGPXV 12
Db 208 DSCRGDSGGPVP 219
RESULT 11
ID Q8MLC4 PRELIMINARY; PRT; 277 AA.
AC Q8MLC4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG30088-PA.
GN ORFNames=CG30088;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,

RA Bettencourt B.R., Ceiniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AF003809; AAM70959.2; -;
DR HSSP; P00750; 1RTP.
DR FlyBase; FBgn0050088; CG30088.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 277 AA; 30725 MW; 8DF1E0C08C789C8 CRC64;

Query Match 93.0%; Score 53; DB 2; Length 277;
Best Local Similarity 75.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCGDGGGPXV 12
DB 217 DTCSGDSGGPLV 228

RESULT 12
Q9D4I3 PRELIMINARY; PRT; 282 AA.
AC Q9D4I3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE library, clone:4931440B9 product:TESTIS SERINE PROTEASE-1
DE homolog.
DE Name=4931440B09Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Carninci P., Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690 (2001).
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA The FANTOM Consortium;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Iehikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AK016509; BAB30277.1; -;
DR HSSP; P00766; 1CHG.
DR MEROPS; S01.417; -;
DR MGD; MGI:1918253; 4931440B09Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 282 AA; 31997 MW; 48D9B74F750BAB4 CRC64;

Query Match 93.0%; Score 53; DB 2; Length 282;
Best Local Similarity 75.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCGDGGGPXV 12
DB 200 DTCSGDSGGPLV 211

RESULT 13

Q8T9U6 Q8T9U6 PRELIMINARY; PRT; 296 AA.
 AC Q8T9U6;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative serine protease.
 OS Aedes aegypti (yellowfever mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
 OX NCBI_TaxID=7159;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Black eye; TISSUE=Salivary gland;
 RX MEDLINE=2202207; PubMed=1231246; DOI=10.1016/S0965-1748(02)00047-4;
 RA Valenzuela J.G., Pham V.M., Garfield M.K., Francischetti I.M.,
 RA Ribeiro J.M.;
 RT "Toward a description of the salivary gland of the adult female mosquito
 RT Aedes aegypti";
 RL Insect Biochem. Mol. Biol. 32:1101-1122(2002).
 CC -!- SIMILARITY: Belongs to peptidase family S1.
 DR EMBL; AF466600; AAL76023.1; -;
 DR HSP; P07338; IKDO.
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004295; F:trypsin activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001254; Peptidase S1.
 DR InterPro; IPR001314; Peptidase S1.
 DR InterPro; IPR009003; Pept_Ser_Cys.
 DR Pfam; PF00089; Trypsin; I.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00020; TRYSP; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS0134; TRYPSIN_HIS; UNKNOWN_1.
 KW Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 296 AA; 32221 MW; CESD99363F9363133 CRC64;
 Query Match 93.0%; Score 53; DB 2; Length 296;
 Best Local Similarity 75.0%; Pred. NO. 0.21;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 DXCGDSDGSGPV 12
 DB 243 DSCDSDGSGGPAV 254
 RESULT 14
 Q9VS87 Q9VS87 PRELIMINARY; PRT; 299 AA.
 AC Q9VS87;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE CG32374-PA.
 GN ORFNames=CG32374;
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Flosser C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skujski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tuip J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to peptidase family S1.
 DR EMBL; AE003558; AAF50540.2; -;

DR HSP; P00760; LEZX.
DR FlyBase; FBgn0052374; CG32374.
DR GO; GO:0004295; F:trypsin activity; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1a.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 299 AA; 33592 MW; 2C56844DAB3227D8 CRC64;

Query Match 93.0%; Score 53; DB 2; Length 299;
Best Local Similarity 75.0%; Pred. NO. 0.21;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
Db 247 DTCSGDSGGPLV 258

RESULT 15

Q920S2 PRELIMINARY; PRT; 322 AA.
AC Q920S2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Testis serine protease-1.
GN Name=4931440B09Rik; Synonyms=tessp-1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Matsui H., Takano N., Takahashi T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AB049453; BAB68561.1; -.
DR HSP; P00766; 1CHG.
DR MEROPS; S01.417; -.
DR MGD; MGI:1918253; 4931440B09Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 322 AA; 36218 MW; ED55F9A199A3E491 CRC64;

Query Match 93.0%; Score 53; DB 2; Length 322;
Best Local Similarity 75.0%; Pred. NO. 0.23;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
Db 240 DTCSGDSGGPLV 251

Search completed: November 1, 2005, 23:37:59
Job time : 49.2941 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2005, 23:29:56 ; Search time 10.3529 Seconds
(without alignments)
111.524 Million cell updates/sec

Title: US-10-766-752B-9
Perfect score: 57 DXCXGDSGGPXV 12
Sequence: 1 DXCXGDSGGPXV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: piri:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	93.0	251	2	29K serine proteinase
2	53	93.0	699	1	Ra-reactive factor
3	53	93.0	1019	2	coagulation factor
4	52	91.2	66	2	kallikrein - mouse
5	52	91.2	149	1	tissue kallikrein
6	52	91.2	161	2	coagulation factor
7	52	91.2	161	2	coagulation factor
8	52	91.2	191	2	complement factor
9	52	91.2	229	1	trypsin (EC 3.4.21
10	52	91.2	229	1	trypsin (EC 3.4.21
11	52	91.2	231	1	TRPGTR
12	52	91.2	231	2	trypsin (EC 3.4.21
13	52	91.2	234	2	F42696
14	52	91.2	235	2	thrombin (EC 3.4.2
15	52	91.2	235	2	thrombin (EC 3.4.2
16	52	91.2	236	2	thrombin (EC 3.4.2
17	52	91.2	236	2	thrombin (EC 3.4.2
18	52	91.2	238	1	TRWV5Y
19	52	91.2	238	2	S31779
20	52	91.2	239	2	G42696
21	52	91.2	240	2	S39047
22	52	91.2	241	2	S39048
23	52	91.2	242	2	S31776
24	52	91.2	242	2	S49489
25	52	91.2	242	2	trypsin (EC 3.4.21
26	52	91.2	243	2	A35871
27	52	91.2	245	2	kallikrein-like se
28	52	91.2	245	2	complement factor
29	52	91.2	246	1	trypsin (EC 3.4.21

ALIGNMENTS

RESULT 1

PC1235
29K serine proteinase (EC 3.4.21.-) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-May-2004
C:Accession: PC1235; PC1236
R:Takahashi, A.; Takayama, Y.; Hattuse, H.; Kawakami, M.
Biochem. Biophys. Res. Commun. 190, 681-687, 1993
A:Title: Presence of a serine protease in the complement-activating component of the co
A:Reference number: PC1235; MUID:93176166; PMID:8439319
A:Accession: PC1235
A:Molecule type: mRNA
A:Residues: 1-251 <TAK>
A:Accession: PC1236
A:Molecule type: Protein
A:Residues: 1-18;12-135;155-173;182-201 <TA2>
C:Superfamily: complement-activating serine proteases Cir/Cis/WASP; Cir/Cis repeat homo
C:Keywords: hydrolase; serine proteinase
F:1-243/Domain: trypsin homology <TRY>

Query Match 93.0%; Score 53; DB 2; Length 251;
Best Local Similarity 75.0%; Pred. No. 0.043;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | | | |
DB 192 DACAGDSGGPMV 203

RESULT 2

I54763
Ra-reactive factor (EC 3.4.21.-) 1 precursor - human
N:Alternate names: mannose binding protein-associated serine proteinase 1 (MASP-1)
C:Species: Homo sapiens (man)
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: I54763; JN0883
R:Sato, T.; Endo, Y.; Matsushita, M.; Fujita, T.
Int. Immunol. 6, 665-669, 1994
A:Title: Molecular characterization of a novel serine protease involved in activation o
A:Reference number: I54763; MUID:94289349; PMID:8018603
A:Accession: I54763
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-699 <SAT>
A:Cross-references: UNIPROT:P48740; GB:D28593; NID:G790963; PIDN:BA05928.1; PID:G47112
R:Takada, F.; Takayama, Y.; Hattuse, H.; Kawakami, M.
Biochem. Biophys. Res. Commun. 196, 1003-1009, 1993
A:Title: A new member of the C1s family of complement proteins found in a bactericidal
A:Reference number: JN0883; MUID:94059062; PMID:8240317
A:Accession: JN0883
A:Molecule type: mRNA
A:Residues: 1-234, 'S', 236-284, 'G', 286-498, 'K', 500-542, 'S', 544-642, 'S', 644-699 <TAK>

A:Cross-references: DDBJ:D17525; NID:g439712; PIDN:BAA04477.1; PID:g439713
 A:Experimental source: liver
 C:Comment: this is a serum bactericidal factor that activates complement C4 and C2 components
 C:Genetics:
 A:Gene: GDB:MSP1; GDB:CRARF; CRARF1; PRS85; MASP
 A:Cross-references: GDB:361104; GDB:330954; OMIM:600521
 A:Map position: 3q27-3q28
 C:Superfamily: complement-activating serine proteases C1r/C1s/MASP; C1r/C1s repeat homolog
 C:Keywords: beta-hydroxyasparagine; complement pathway; duplication; glycoprotein; hydro
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-448/449/Product: Ra-reactive factor #status predicted <MAT>
 F:19-135/Domain: C1r/C1s repeat homology <C1R1>
 F:143-181/Domain: EGF homology <EGF>
 F:185-294/Domain: C1r/C1s repeat homology <C1R2>
 F:301-362/Domain: complement factor H repeat homology <PH1>
 F:367-432/Domain: complement factor H repeat homology <PH2>
 F:449-691/Domain: trypsin homology <TRY>
 F:49-178/407/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F:73-91/143-157/153-166/168-181/185-212/242-260/301-349/329-362/367-414/397-432/436-572,
 F:159/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
 F:448-449/Cleavage site: Arg-Ile (autolytic) #status predicted
 F:490/552/646/Active site: His, Asp, Ser #status predicted

Query Match 93.0%; Score 53; DB 1; Length 699;
 Best Local Similarity 75.0%; Pred. No. 0.11;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXPV 12
 | | | | | | | | | |
 Db 640 DACAGDSGGPMV 651

RESULT 3
 A38738
 coagulation factor C precursor - horseshoe crab (Tachyplesus tridentatus)
 N:Alternate names: coagulation-complement factor C; Limulus factor C
 N:Contains: coagulation factor C heavy chain; coagulation factor C light chain peptide A
 C:Species: Tachyplesus tridentatus
 C:Date: 04-Oct-1991 #sequence_revision 04-Oct-1991 #text_change 09-Jul-2004
 C:Accession: A38738; B38738; S00105
 R:Muta, T.; Miyata, T.; Misumi, Y.; Tokunaga, F.; Nakamura, T.; Toh, Y.; Ikehara, Y.; Iw
 J. Biol. Chem. 266, 6554-6561, 1991
 A:Title: Limulus factor C. An endotoxin-sensitive serine protease zymogen with a mosaic
 A:Reference number: A38738; MUID:91177916; PMID:2007602
 A:Accession: A38738
 A:Molecule type: mRNA
 A:Residues: 1-1019 <MUT>
 A:Cross-references: UNIPROT:P28175; GB:D90271; NID:g217396; PID:d1015020; PID:g217397
 A:Accession: B38738
 A:Molecule type: mRNA
 A:Residues: 1-466/616, 'DN', 619-620, 'A', 622 <MUT>
 A:Cross-references: GB:D90272
 R:Tokunaga, F.; Miyata, T.; Nakamura, T.; Morita, T.; Kuma, K.I.; Miyata, T.; Iwanaga, S
 Eur. J. Biochem. 167, 405-416, 1987
 A:Title: Lipopolysaccharide-sensitive serine-protease zymogen (factor C) of horseshoe cr
 is a novel type of serine protease.
 A:Reference number: S00105; MUID:88004461; PMID:3308457
 A:Accession: S00105
 A:Molecule type: protein
 A:Residues: 'S', 27-39, 'S', 41, 'T', 43-54, 'E', 691-782; 950-977 <TOK>
 A:Superfamily: coagulation factor C; C-type lectin homology; complement factor H repeat
 C:Keywords: alternative splicing; glycoprotein; hemolymph coagulation; hydrolase; serine
 F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-690/Product: coagulation factor C heavy chain #status experimental <HCH>
 F:136-195/Domain: complement factor H repeat homology <PH01>
 F:199-254/Domain: complement factor H repeat homology <PH02>
 F:260-321/Domain: complement factor H repeat homology <PH03>
 F:436-564/Domain: C-type lectin homology <LCH>
 F:576-634/Domain: complement factor H repeat homology <PH04>
 F:685-747/Domain: complement factor H repeat homology #status atypical <PH05>
 F:691-762/Product: coagulation factor C light chain peptide A #status experimental <PPA>
 F:763-1019/Product: coagulation factor C light chain peptide B #status experimental <PPB>
 F:763-1015/Domain: trypsin homology <TRY>

F:523, 534, 624, 912/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F:767/Binding site: carboxydrate (Asn) (covalent) #status experimental
 F:809, 865, 966/Active site: His, Asp, Ser #status predicted

Query Match 93.0%; Score 53; DB 2; Length 1019;
 Best Local Similarity 75.0%; Pred. No. 0.16;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXPV 12
 | | | | | | | | | |
 Db 960 DACSGDSGGPLV 971

RESULT 4
 I52972
 kallikrein - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
 C:Accession: I52972
 R:Nordeen, S.K.; Mason, A.J.; Richards, R.I.; Baxter, J.D.; Shine, J.
 DNA 1, 309-311, 1982
 A:Title: Mouse kallikrein arginyl-esteropeptidase genes: Analysis of cloned cDNAs suggest
 A:Reference number: I52972; MUID:83182015; PMID:6926406
 A:Accession: I52972
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-66 <RES>
 A:Cross-references: UNIPROT:Q61752; GB:K01654; NID:g198490; PIDN:AAA39346.1; PID:g198491
 C:Superfamily: trypsin; trypsin homology (fragment)
 F:1-58/Domain: trypsin homology (fragment) <TRY>

Query Match 91.2%; Score 52; DB 2; Length 66;
 Best Local Similarity 66.7%; Pred. No. 0.019;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXPV 12
 | | | | | | | | | |
 Db 12 DTCAGDSGGPLI 23

RESULT 5
 KQMSM
 tissue kallikrein (EC 3.4.21.35), submandibular - mouse (fragment)
 N:Alternate names: glandular kallikrein; kininogenin
 C:Species: Mus musculus (house mouse)
 C:Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 09-Jul-2004
 C:Accession: A00939
 R:Richards, R.I.; Catanzaro, D.F.; Mason, A.J.; Morris, B.J.; Baxter, J.D.; Shine, J.
 J. Biol. Chem. 257, 2758-2761, 1982
 A:Title: Mouse glandular kallikrein genes. Nucleotide sequence of cloned cDNA coding for
 A:Reference number: A00939; MUID:82142394; PMID:6174512
 A:Accession: A00939
 A:Molecule type: mRNA
 A:Residues: 1-149 <RIC>
 A:Cross-references: UNIPROT:Q925V8; GB:V00828; GB:J00389; NID:g52773; PIDN:CAA24211.1; I
 A:Experimental source: Ouakenbush inbred strain
 C:Comment: Tissue kallikreins cleave Met-Lys and Arg-Ser bonds in kininogen to release 1
 C:Genetics:
 A:Map position: 7
 A:Superfamily: trypsin; trypsin homology
 C:Keywords: hydrolase; saliva; serine proteinase; submandibular gland
 F:1-141/Domain: trypsin homology (fragment) <TRY>
 F:8,101/Active site: Asp, Ser #status predicted
 F:40-107/72-86, 97-122/Disulfide bonds: #status predicted

Query Match 91.2%; Score 52; DB 1; Length 149;
 Best Local Similarity 66.7%; Pred. No. 0.04;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXPV 12
 | | | | | | | | | |
 Db 95 DTCAGDSGGPLI 106


```
RESULT 6
I62744
coagulation factor Xa (EC 3.4.21.6) - rhesus macaque (fragment)
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C:Accession: I62744
R:Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Harada, M.; Niho, Y.
Eur. J. Haematol. 52, 162-168, 1994
A:Title: Analysis of the partial nucleotide sequences and deduced primary structures of
A:Reference number: I46196; MUID:94222160; PMID:8168596
A:Accession: I62744
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-161 <RES>
A:Cross-references: UNIPROT:Q28511; GB:D21214; NID:G415307; PIDN:BAA04755.1; PID:G455393
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxylglutam
F:1-161/Domain: trypsin homology (fragment) <TRY>

Query Match          91.2%; Score 52; DB 2; Length 161;
Best Local Similarity 75.0%; Pred. No. 0.043;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXY 12
| | | | | | | |
DB 117 DACQGDGGPHV 128

RESULT 7
I48158
coagulation factor Xa (EC 3.4.21.6) - golden hamster (fragment)
C:Species: Mesocricetus auratus (golden hamster)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I48158
R:Murakawa, M.; Okamura, T.; Kamura, T.; Kuroiwa, M.; Harada, M.; Niho, Y.
Eur. J. Haematol. 52, 162-168, 1994
A:Title: Analysis of the partial nucleotide sequences and deduced primary structures of
A:Reference number: I46196; MUID:94222160; PMID:8168596
A:Accession: I48158
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-161 <RES>
A:Cross-references: UNIPROT:Q60546; GB:D21216; NID:G415304; PIDN:BAA04757.1; PID:G455393
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology
C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxylglutam
F:1-161/Domain: trypsin homology (fragment) <TRY>

Query Match          91.2%; Score 52; DB 2; Length 161;
Best Local Similarity 75.0%; Pred. No. 0.043;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXY 12
| | | | | | | |
DB 117 DACQGDGGPHV 128

RESULT 8
S54115
complement factor D (EC 3.4.21.46) - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C:Accession: S54115
R:Nicolas, N.
submitted to the EMBL Data Library, April 1995
A:Reference number: S54115
A:Accession: S54115
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-191 <NIC>
A:Cross-references: UNIPROT:P51779; EMBL:Z49058; NID:G773264; PIDN:CAA88844.1; PID:G7732
C:Superfamily: trypsin; trypsin homology
C:Keywords: complement alternate pathway; hydrolase; plasma; serine proteinase
```

```
F:1-181/Domain: trypsin homology (fragment) <TRY>

Query Match          91.2%; Score 52; DB 2; Length 191;
Best Local Similarity 75.0%; Pred. No. 0.05;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXY 12
| | | | | | | |
DB 135 DSCCKGDSGGPLV 146

RESULT 9
TRB0TR
trypsin (EC 3.4.21.4) precursor - bovine
N:Contains: trypsinogen
C:Species: Bos primigenius taurus (cattle)
C>Date: 24-Apr-1984 #sequence_revision 28-Feb-1986 #text_change 18-Jul-1997
C:Accession: A90164; A0946; S08774
R:Mikes, O.; Holeysovsky, V.; Tomasek, V.; Sorm, F.
Biochem. Biophys. Res. Commun. 24, 346-352, 1966
A:Title: Covalent structure of bovine trypsinogen. The position of the remaining amides
A:Reference number: A90164; MUID:67168848; PMID:5967094
A:Accession: A90164
A:Molecule type: protein
A:Residues: 1-57; 'Q', 'S', '67', 'Q', '69-150', 'N', '152-176', 'N', '178-229' <MIK>
R:Hartley, B.S.
Philos. Trans. R. Soc. Lond. B257, 77-87, 1970
A:Reference number: A93755
A:Content: annotation; revisions
R:Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 14, 1358-1366, 1975
A:Title: Amino acid sequence of dogfish trypsin.
A:Reference number: A00950; MUID:75146445; PMID:1092332
A:Content: annotation; revisions
A:Note: the sequence agrees with that shown
R:Bode, W.; Schwager, P.
J. Mol. Biol. 98, 693-717, 1975
A:Title: The refined crystal structure of bovine beta-trypsin at 1.8 angstrom resolution
A:Reference number: A92954; MUID:76072097; PMID:512
A:Content: annotation; X-ray crystallography; binding sites for calcium, substrate, an
C:Comment: Trypsinogen is synthesized in the acinar cells of the pancreas,
C:Comment: Autocatalytic cleavage after Lys-6 leads to beta-trypsin by releasing a term
s pseudotrypsin. A cleavage may also occur after Arg-105.
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-229/Product: trypsinogen #status experimental <ZYM>
F:1-6/Domain: activation peptide #status experimental <APT>
F:7-222/Domain: trypsin homology <TRY>
F:7-131,132-229/Product: alpha-trypsin #status experimental <MPT>
F:6-7/Cleavage site: Lys-11e (enteropeptidase) #status experimental
F:13-143,31-47,115-216,122-189,154-168,179-203/Disulfide bonds: #status experimental
F:46,90,183/Active site: His, Asp, Ser #status experimental
F:58,60,63,68/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental
F:131-132/Cleavage site: Lys-Ser (autolytic) #status experimental

Query Match          91.2%; Score 52; DB 1; Length 229;
Best Local Similarity 75.0%; Pred. No. 0.06;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXY 12
| | | | | | | |
DB 177 DSCQGDGGPXY 188

RESULT 10
TRDFS
trypsin (EC 3.4.21.4) precursor - spiny dogfish
N:Alternate names: trypsinogen
C:Species: Squalus acanthias (spiny dogfish)
C>Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 09-Jul-2004
C:Accession: A00950; B27719
R:Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 14, 1358-1366, 1975
```

A>Title: Amino acid sequence of dogfish trypsin.
A:Reference number: A00950; MUID:75146445; PMID:1092332

A:Accession: A00950

A:Molecule type: protein

A:Residues: 8-229 <TIT>

A:Cross-references: UNIPROT:P00764

A>Note: 119-Pro was also found

R:Hermanson, M.A.; Tye, R.W.; Reeck, G.R.; Neurath, H.; Walsh, K.A.

PNBS Lett. 14, 222-224, 1971

A>Title: Comparison of the amino terminal sequences of bovine, dogfish, and lungfish try

A:Reference number: A27719

A:Accession: B27719

A:Molecule type: protein

A:Residues: 1-21 <HER>

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase; zym

F:1-7/Domain: activation peptide #status experimental <APT>

F:8-229/Product: trypsin #status predicted <MAT>

F:8-222/Domain: trypsin homology <TRY>

F:14-143,32-48,116-216,123-189,154-168,179-203/Disulfide bonds: #status predicted

F:47,91,183/Active site: His, Asp, Ser #status predicted

F:159,61,64,69/Binding site: calcium (Glu, Asp, Ala, Glu) #status predicted

Query Match 91.2%; Score 52; DB 1; Length 229;

Best Local Similarity 75.0%; Pred. No. 0.06;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12

| | | | | | | |

Db 177 DSCQDGGGPV 188

RESULT 11

TRPGTR

trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)

N:Contains: trypsinogen

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 24-Apr-1984 #sequence revision 24-Apr-1984 #text_change 09-Jul-2004

C:Accession: A90641; A90368; A00947

R:Charles, M.; Roversy, M.; Guidoni, A.; Desnuelle, P.

Biochim. Biophys. Acta 69, 115-129, 1963

A>Title: Su le trypsinegene et la trypsine de porc.

A:Reference number: A90641

A:Accession: A90641

A:Molecule type: protein

A:Residues: 1-10 <CHA>

A:Cross-references: UNIPROT:P00761

R:Hermanson, M.A.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.

Biochemistry 12, 3146-3153, 1973

A>Title: Determination of the amino acid sequence of porcine trypsin by sequenator analy

A:Reference number: A90368; MUID:73258692; PMID:4738933

A:Accession: A90368

A:Molecule type: protein

A:Residues: 9-231 <HER>

A>Note: at position 20, Ile and Val occur alternatively

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase; zym

F:1-231/Product: trypsinogen #status experimental <ZIFM>

F:1-8/Domain: activation peptide #status experimental <APT>

F:9-231/Product: trypsin #status experimental <MAT>

F:9-224/Domain: trypsin homology <TRY>

F:15-145,33-49,117-218,124-191,156-170,181-205/Disulfide bonds: #status predicted

F:48,92,185/Active site: His, Asp, Ser #status predicted

F:160,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 91.2%; Score 52; DB 1; Length 231;

Best Local Similarity 75.0%; Pred. No. 0.06;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12

| | | | | | | |

Db 179 DSCQDGGGPV 190

RESULT 12

S31778

trypsin (EC 3.4.21.4) II precursor - Atlantic salmon (fragment)

C:Species: Salmo salar (Atlantic salmon)

C>Date: 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C:Accession: S66658; S31778

R:Male, R.; Lorens, J.B.; Snaas, A.O.; Torrisen, K.R.

Eur. J. Biochem. 232, 677-685, 1995

A>Title: Molecular cloning and characterization of anionic and cationic variants of tryp

A:Reference number: S66657; MUID:96035908; PMID:7556223

A:Accession: S66658

A:Molecule type: mRNA

A:Residues: 1-231 <MAL>

A:Cross-references: UNIPROT:P35032; EMBL:X70073; NID:G64385; PIDN:CAA49678.1; PID:G64386

C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase

F:1-4/Domain: signal sequence (fragment) #status predicted <SIG>

F:5-9/Domain: activation peptide #status predicted <APT>

F:10-231/Product: trypsin II #status predicted <MAT>

F:10-224/Domain: trypsin homology <TRY>

F:16-145,34-50,118-218,125-191,156-170,181-205/Disulfide bonds: #status predicted

F:49,93,185/Active site: His, Asp, Ser #status predicted

Query Match 91.2%; Score 52; DB 2; Length 231;

Best Local Similarity 75.0%; Pred. No. 0.06;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12

| | | | | | | |

Db 179 DSCQDGGGPV 190

RESULT 13

F42696

thrombin (EC 3.4.21.5) B chain - Cynops pyrogastor (fire-bellied newt) (fragment)

C:Species: Cynops pyrogastor (fire-bellied newt)

C>Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004

C:Accession: F42696

R:Banfield, D.K.; MacGillivray, R.T.A.

Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992

A>Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq

A:Reference number: A42696; MUID:92212913; PMID:1557383

A>Note: sequence not

A:Accession: F42696

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra

A:Molecule type: mRNA

A:Residues: 1-234 <BAN>

A:Cross-references: UNIPROT:Q90387; GB:M81395

C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology

C:Keywords: hydrolase; serine proteinase

Query Match

Best Local Similarity 91.2%; Score 52; DB 2; Length 234;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12

| | | | | | | |

Db 174 DACEGDSGGPFV 185

RESULT 14

D42696

thrombin (EC 3.4.21.5) B chain - chicken (fragment)

C:Species: Gallus gallus (chicken)

C>Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004

C:Accession: D42696

R:Banfield, D.K.; MacGillivray, R.T.A.

Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992

A>Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and seq

A:Reference number: A42696; MUID:92212913; PMID:1557383

A:Accession: D42696

A>Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-235 <BAN>
A:Cross-references: UNIPROT:Q91001; GB:M81391
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 91.2%; Score 52; DB 2; Length 235;
Best Local Similarity 75.0%; Pred. No. 0.061;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | | | |
Db 175 DACEGDSGGPFV 186

RESULT 15
E42696
thrombin (EC 3.4.21.5) B chain - tokay (fragment)
C:Species: Gekko gecko (tokay)
C>Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C:Accession: E42696
R:Banfield, D.K.; MacGillivray, R.T.A.
Proc. Natl. Acad. Sci. U.S.A. 89, 2779-2783, 1992
A:Title: Partial characterization of vertebrate prothrombin cDNAs: amplification and sequencing
A:Reference number: A42696; MUID:92212913; PMID:1557383
A:Accession: E42696
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-235 <BAN>
A:Cross-references: UNIPROT:Q91004; GB:M81392
C:Superfamily: thrombin; Gla domain homology; kringle homology; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-226/Domain: trypsin homology (fragment) <TRY>

Query Match 91.2%; Score 52; DB 2; Length 235;
Best Local Similarity 75.0%; Pred. No. 0.061;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | | | |
Db 175 DACEGDSGGPFV 186

Search completed: November 1, 2005, 23:43:08
Job time : 11.3529 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 1, 2005, 23:25:16 ; Search time 60.2353 Seconds
(without alignments)
77.050 Million cell updates/sec

Title: US-10-766-752B-9

Perfect score: 57
Sequence: 1 DXCXGSGGPXV 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 segs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53	93.0	248	4	ABB65295 Drosophil
2	53	93.0	252	4	ABB63479 Drosophil
3	53	93.0	264	8	ADQ08746 Ciona int
4	53	93.0	296	4	AAB85076 Amino aci
5	53	93.0	346	3	AAY82701 Tick deri
6	53	93.0	362	4	ABB62017 Drosophil
7	53	93.0	477	4	ABB65796 Drosophil
8	53	93.0	580	4	ABB60547 Drosophil
9	53	93.0	611	4	ABB65794 Drosophil
10	53	93.0	623	8	ADP04651 Sea squir
11	53	93.0	698	8	ABM83722 Human dia
12	53	93.0	699	7	ADE87459 Human MBL
13	53	93.0	699	8	ADL91028 Human man
14	53	93.0	774	4	ABB70674 Drosophil
15	53	93.0	1019	2	Aaw43394 Singapore
16	53	93.0	1019	2	AAY05750 Drosophil
17	53	93.0	1019	2	AAY94302 Drosophil
18	53	93.0	1019	3	AAY42490 Drosophil
19	53	93.0	1019	6	ABP72332 Drosophil
20	53	93.0	1019	6	ABP72332 Drosophil
21	53	93.0	1019	6	ABP72332 Drosophil
22	53	93.0	1083	2	Aaw43393 Singapore
23	53	93.0	1083	2	AAY05749 Drosophil
24	53	93.0	1083	2	Aaw94301 Drosophil
25	53	93.0	1083	3	AAY42489 Drosophil

26	53	93.0	1083	4	AAB60934	Aab60934 Horsehoe
27	53	93.0	1083	6	ABP72333	Abp72333 Horsehoe
28	53	93.0	1190	4	ABB62480	Abb62480 Drosophil
29	52	91.2	12	3	AAY72111	Aay72111 Peptide f
30	52	91.2	12	3	AAY72112	Aay72112 Peptide f
31	52	91.2	12	4	AAB71676	Aab71676 Human col
32	52	91.2	12	5	ABG30800	Abg30800 Human ser
33	52	91.2	12	5	ABG31651	Abg31651 Human ser
34	52	91.2	12	5	AAE17931	Aae17931 Human gen
35	52	91.2	12	5	AAE17932	Aae17932 Human gen
36	52	91.2	12	5	AAM50857	Aam50857 Serine es
37	52	91.2	12	6	ADA44894	Ada44894 Human ser
38	52	91.2	12	7	ADF17980	Adf17980 Serine es
39	52	91.2	13	2	AAY31238	Aay31238 Rat Facto
40	52	91.2	14	5	AAE17240	Aae17240 Human tra
41	52	91.2	14	5	AAE18999	Aae18999 Human mat
42	52	91.2	14	5	AAU78375	Aau78375 Thrombin
43	52	91.2	14	5	AAE20158	Aae20158 Human thr
44	52	91.2	14	7	ABB80264	Abb80264 Thrombin
45	52	91.2	14	7	ADF17985	Adf17985 Human pro

ALIGNMENTS

RESULT 1
ABB65295
ID ABB65295 standard; protein; 248 AA.
XX AC ABB65295;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 22677.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers BW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL09398.
XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.
XX Disclosure; SEQ ID NO 22677; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX 30-NOV-2000; 2000WO-DK000659.
 XX
 XX 02-DEC-1999; 99DK-00001721.
 PR 21-JUL-2000; 2000DK-00001126.
 XX
 XX (JENS/) JENSENIUS J C.
 PA (THIE/) THIEL S.
 XX
 XX Jensenius JC, Thiel S;
 PI
 XX WPI; 2001-374820/39.
 DR
 XX
 XX Novel pure mannan-binding lectin associated serine protease polypeptides
 PT and polynucleotides encoding the protein, used for treating reoxygenated
 PT ischemic tissues, mannon-binding lectin deficiency, multiple sclerosis.
 XX
 XX Disclosure; Fig 6; 99pp; English.
 PS
 XX The invention relates to a mannan-binding lectin (MBL) associated serine
 CC protease-3 (MASP-3) polypeptide. MASP-3 exerts an inhibitory effect on
 CC the complement activation particularly when bound to MBL/MASP-2 complexes
 CC and directly activates complement system through binding to MBL. The MASP
 CC -3 polypeptides, polynucleotides and modulators are useful for preparing
 CC a pharmaceutical composition for treating aberrant MASP-3 activity such
 CC as infections, cancer, MBL-deficiency, disorders of the immune system and
 CC reproductive system. The MASP-3 polypeptide is also used for treating
 CC diseases associated with human immunodeficiency virus, multiple
 CC sclerosis, myasthenia gravis, epilepsy, rheumatoid arthritis, vasculitis,
 CC autoimmune hemolytic anemia, Crohn's disease, asthma, diabetes,
 CC psoriasis, multiple myeloma, atherosclerosis etc. It is also useful for
 CC inhibiting activation of C4 complement by inhibiting the MBL pathway,
 CC inhibiting MASP-2 activity, inhibiting or treating an inflammatory
 CC condition related to complement activation through MBL/MASP complexes.
 CC MASP-3 is also useful for treating an inflammatory condition resulting
 CC from an autoimmune condition after acute myocardial infarction or brain
 CC ischemia. It is also useful for treating an individual suffering from a
 CC disorder resulting from an imbalanced cytokine network. The present
 CC sequence represents a MASP-1 polypeptide, used for comparison studies
 CC with MASP-3
 XX
 XX Sequence 296 AA;
 SQ

Query Match 93.0%; Score 53; DB 4; Length 296;
 Best Local Similarity 75.0%; Pred. No. 3.7;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
 DB 237 DACAGDSGGPMV 248
 |||||
 |||||

RESULT 5
 AAY82701
 ID AAY82701 standard; protein; 346 AA.
 XX
 XX AAY82701;
 AC
 XX 10-AUG-2000 (first entry)
 DT
 XX Tick derived serine protease protein sequence #2.
 DE
 XX Tick; vaccine; infection; salivary gland antigen; immunogen;
 KW serine protease; cysteine protease; blood sucker.
 XX
 XX Haemaphysalis longicornis.
 OS
 XX JP2000083677-A.
 PN
 XX 28-MAR-2000.
 PD
 XX 17-SEP-1998; 98JP-00281932.
 PF
 XX

17-SEP-1998; 98JP-00281932.
 PR (FARB) BAYER KK.
 XX
 XX WPI; 2000-296340/26.
 DR N-PSDB; AAA29621.
 DR
 XX A gene encoding tick salivary gland antigen - useful as a vaccine for
 PT protecting animals from tick-carried infections.
 PT
 XX Claim 7; Page 13-14; 29pp; Japanese.
 PS
 XX The present sequence represents a tick derived serine protease. The
 CC present invention also describes a tick salivary gland antigen related
 CC immunogen and a tick derived cysteine protease. A nucleotide sequence
 CC encoding any of the above proteins can be used in a vaccine against tick
 CC carried infections for domestic animals such as cattle
 XX
 XX Sequence 346 AA;
 SQ

Query Match 93.0%; Score 53; DB 3; Length 346;
 Best Local Similarity 75.0%; Pred. No. 4.2;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
 DB 287 DACSGDSGGPMV 298
 |||||
 |||||

RESULT 6
 ABB62017
 ID ABB62017 standard; protein; 362 AA.
 XX
 XX ABB62017;
 AC
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster polypeptide SEQ ID NO 12843.
 DE
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US009231.
 PF
 XX 23-MAR-2000; 2000US-0191637P.
 PR
 XX 11-JUL-2000; 2000US-00614150.
 PR
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL06120.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 PT
 XX Disclosure; SEQ ID NO 12843; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB577737-

CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 362 AA;

Query Match 93.0%; Score 53; DB 4; Length 362;
Best Local Similarity 75.0%; Pred. No. 4.4;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
| | | | | | | |
Db 239 DACAGDSGGPLV 250

RESULT 7

ABB65796
ID ABB65796 standard; protein; 477 AA.

XX AC ABB65796;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 24180.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL09899.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions.

XX PS Disclosure; SEQ ID NO 24180; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL01840-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
ABBS72072). The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 477 AA;

Query Match 93.0%; Score 53; DB 4; Length 477;
Best Local Similarity 75.0%; Pred. No. 5.6;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12

Db 201 DSCGDSGGPLV 212

RESULT 8

ABB60547
ID ABB60547 standard; protein; 580 AA.

XX AC ABB60547;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 8433.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL04650.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions.

XX PS Disclosure; SEQ ID NO 8433; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
ABBS72072). The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 580 AA;

Query Match 93.0%; Score 53; DB 4; Length 580;
Best Local Similarity 75.0%; Pred. No. 6.6;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12

Db 528 DTCSGDSGGPLV 539

RESULT 9

ABB65794
ID ABB65794 standard; protein; 611 AA.

XX AC ABB65794;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 24174.

XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.

OS *Drosophila melanogaster*.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI N-PSDB; ABL09897.
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL09897.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 24174; 2lpp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (AB557737-
 CC AB572072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 611 AA;
 SQ
 Query Match 93.0%; Score 53; DB 4; Length 611;
 Best Local Similarity 75.0%; Pred. No. 7;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 DXCXGDSGGPXV 12
 DB 203 DACAGDSGGPVV 214
 RESULT 10
 ADP04651
 ID ADP04651 standard; protein; 623 AA.
 XX AC ADP04651;
 XX 29-JUL-2004 (first entry)
 XX Sea squirt protein with tissue specific expression in development Seq246.
 DE sea squirt; regeneration medicine; gene therapy; cell proliferation;
 XX differentiation; reproduction; environmental measurement; water survey.
 KW Ciona intestinalis.
 XX JP2004057129-A.
 XX 26-FEB-2004.
 XX 31-JUL-2002; 2002JP-00222593.
 XX 31-JUL-2002; 2002JP-00222593.
 XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 PA WPI; 2004-287079/27.
 DR N-PSDB; ADP04650.

XX Novel gene cluster which is specifically expressed in tissue or organ
 PT during developmental phase of sea squirt, useful for elucidation of
 PT mechanism of development of tissue or organ of sea squirt.
 XX Claim 1; SEQ ID NO 246; 1846pp; Japanese.
 XX This invention relates to novel genes and the encoded proteins thereof
 CC that are derived from the sea squirt *Ciona intestinalis*. Specifically, it
 CC refers to those genes that are expressed in the tissues or organs of the
 CC sea squirt during its developmental phase. The present invention
 CC describes the identification of these genes as useful for elucidation of
 CC the mechanism of development and hence for developing regeneration
 CC medicines and gene therapy techniques. Accordingly, they can be used in
 CC the research of various genetic diseases, as well as the analysis of cell
 CC proliferation, differentiation and reproduction. Furthermore, such
 CC compositions can be useful for environmental measurements and water
 CC surveys, particularly for sea water surveys, and also for the preparation
 CC of transformed sea squirt for improving edibility of sea squirt such as
 CC *Halocynthia roretzi*. This polypeptide sequence is a sea squirt protein
 CC sequence that has tissue specific expression during development, given in
 CC an exemplification of the invention.
 XX Sequence 623 AA;
 SQ
 Query Match 93.0%; Score 53; DB 8; Length 623;
 Best Local Similarity 75.0%; Pred. No. 7.1;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 DXCXGDSGGPXV 12
 DB 551 DTCTGDSGGPLV 562
 RESULT 11
 ABM83722
 ID ABM83722 standard; protein; 698 AA.
 XX AC ABM83722;
 XX 18-NOV-2004 (first entry)
 XX Human diagnostic and therapeutic pprotein SEQ ID NO:3971.
 DE gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
 KW Homo sapiens.
 XX WO2004023973-A2.
 XX 25-MAR-2004.
 XX 12-SEP-2003; 2003WO-US028227.
 XX 12-SEP-2002; 2002US-0410259P.
 PR 12-SEP-2002; 2002US-0410260P.
 XX (INCV-) INCVTE CORP.
 XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
 PI Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
 PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
 PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;
 PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
 PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
 PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
 PI Fatyry S, Shi X, Suarez CJ;
 XX WPI; 2004-329368/30.
 DR N-PSDB; ACN42374.
 XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
 PT in diagnosing a condition, disease or disorder associated with human

PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
 PT in gene mapping.
 XX
 PS Claim 27; Page; 190pp; English.
 XX
 CC The invention relates to novel diagnostic and therapeutic polynucleotides
 CC selected from one of the 2722 sequences defined in the specification. A
 CC polynucleotide of the invention may have a use in gene therapy. The human
 CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
 CC used to diagnose a particular condition, disease or disorder associated
 CC with human molecules, e.g. cell proliferative disorders,
 CC autoimmune/inflammatory disorder, developmental disorder, endocrine
 CC disorder, neurological disorders, gastrointestinal disorders, or
 CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dithp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
 XX
 SQ Sequence 698 AA;

Query Match 93.0%; Score 53; DB 8; Length 698;
 Best Local Similarity 75.0%; Pred. No. 7.8;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
 | | | | | | | |
 DB 639 DACSGDSGGPMV 650

RESULT 12
 ADE87459
 ID ADE87459 standard; protein; 699 AA.
 XX
 AC ADE87459;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human MBL-associated serine protease-1 protein.
 XX
 KW mannan binding lectin; MBL; MBL-associated serine protease; MASP; MASP-1;
 KW MASP-2; MASP-3; antibacterial; immunostimulant; immunocompromised.
 XX
 OS Homo sapiens.
 XX
 PN EP1344533-A1.
 XX
 PD 17-SEP-2003.

XX
 PF 14-MAR-2003; 2003EP-00388016.
 XX
 PR 15-MAR-2002; 2002DK-00000414.
 XX
 PA (NATL-) NATLMUNE AS.
 XX
 PI Larsen JL, Kongerslev L;
 XX
 DR WPI; 2003-758222/72.
 XX
 PT A pharmaceutical composition containing mannan binding lectin or its
 PT variant is useful to treat clinical conditions, particularly infection,
 PT and immunosuppressive conditions.
 XX
 PS Claim 17; SEQ ID NO 2; 44pp; English.

XX The invention relates to a novel pharmaceutical composition comprising
 CC additives and at least 200 micrograms/ml protein containing material
 CC where mannan binding lectin (MBL) and/or MBL variant constitute at least
 CC 35% (w/w) of total protein, or comprising at least 400 micrograms/ml MBL
 CC and/or MBL variant. The protein preferably further comprises one or more

CC MBL-associated serine protease(s) (MASP) or its fragments, particularly
 CC MASP-1, MASP-2 or MASP-3 having respectively the 699, 686 or 728 amino
 CC acid sequence fully defined in the specification. The MBL preferably has
 CC the 248 sequence fully defined in the specification or its fragment and
 CC may be purified naturally occurring human serum MBL or recombinant MBL.
 CC The novel pharmaceutical compositions have antibacterial and
 CC immunostimulant activities. The composition can be used to treat a
 CC clinical condition, particularly an infection, especially where the
 CC individual has a subnormal MBL level. The composition may be used to
 CC treat immunocompromised conditions. This sequence represents the human
 CC MBL-associated serine protease-1 protein of the invention.
 XX

SQ Sequence 699 AA;

Query Match 93.0%; Score 53; DB 7; Length 699;
 Best Local Similarity 75.0%; Pred. No. 7.8;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPXV 12
 | | | | | | | |
 DB 640 DACSGDSGGPMV 651

RESULT 13
 ADL91028
 ID ADL91028 standard; protein; 699 AA.

AC ADL91028;

XX 17-JUN-2004 (first entry)

XX Human mannose binding lectin amino acid sequence SEQ ID NO:14.

XX fusion protein; lectin-complement pathway activating protein; collectin;
 KW antimicrobial; protein therapy; infection; mannose binding lectin.

XX Homo sapiens.

XX WO2004024925-A2.

XX 25-MAR-2004.

XX 10-SEP-2003; 2003WO-DK000585.

XX 10-SEP-2002; 2002DK-00001328.

XX (NATI-) NATIMUNE AS.

XX Kongerslev L, Weilguny D, Matthiesen F;

XX WPI; 2004-270049/25.

XX New fusion protein comprising a first polypeptide sequence derived from a
 PT lectin-complement pathway activating protein, and a second polypeptide
 PT sequence derived from a collectin, useful for treating an infection.

PS Disclosure; SEQ ID NO 14; 127pp; English.

XX The present invention describes a fusion protein comprising a first
 CC polypeptide sequence derived from a lectin-complement pathway activating
 CC protein or a its functional homologue, and a second polypeptide sequence
 CC derived from a collectin or its functional homologue, where the
 CC complement activating protein is not a collectin. Also described: (1) an
 CC isolated nucleic acid encoding the fusion protein; (2) a vector
 CC comprising the nucleic acid; (3) a cell comprising the vector; (4)
 CC treating a clinical condition in an individual by administering the
 CC fusion protein; and (5) a medicament for treating or preventing a
 CC clinical condition in an individual, comprising the fusion protein. The
 CC fusion protein has antimicrobial activity, and can be used in protein
 CC therapy. The fusion protein, method and medicament are useful for
 CC treating a clinical condition, preferably an infection, in an individual
 CC suffering from an increased risk of acquiring an infection, where the
 CC individual has subnormal or normal serum MBL level. The present sequence

CC is used in the exemplification of the present invention.

XX Sequence 699 AA;

Query Match 93.0%; Score 53; DB 8; Length 699;
Best Local Similarity 75.0%; Pred. No. 7.8;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPKV 12
| | | | |
Db 640 DACAGDSGGPMV 651

RESULT 14

ABB70674
ID ABB70674 standard; protein; 774 AA.

XX ABB70674;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 38814.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL14777.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.

XX Disclosure; SEQ ID NO 38814; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 774 AA;

Query Match 93.0%; Score 53; DB 4; Length 774;
Best Local Similarity 75.0%; Pred. No. 8.6;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 DXCXGDSGGPKV 12
| | | | |
Db 527 DSCSGDSGGPMV 538

RESULT 15

AAW43394

ID AAW43394 standard; protein; 1019 AA.

XX AAW43394;

XX 21-MAY-1998 (first entry)

DE Singapore horseshoe crab factor C proenzyme (CrFC 21).

XX Factor C; endotoxin; toxin; assay; Singapore horseshoe crab;
KW purification.

XX Carinoscorpis rotundicauda.

XX Key Location/Qualifiers

FT Peptide 1..24

FT Modified-site /label= Sig_peptide

FT Modified-site 523 /note= "N-glycosylated"

FT Modified-site 534 /note= "N-glycosylated"

FT Modified-site 624 /note= "N-glycosylated"

FT Cleavage-site /note= "N-glycosylated"

FT Cleavage-site 690..691 /note= "cleavage into heavy and light chain intermediates"

FT Modified-site 740 /note= "N-glycosylated"

FT Cleavage-site 762..763 /note= "proteolysis of light chain into A and B chains due to endotoxin activation"

FT Modified-site 767 /note= "N-glycosylated"

FT Active-site 809 /note= "catalytic triad residue"

FT Active-site 865 /note= "catalytic triad residue"

FT Modified-site 912 /note= "N-glycosylated"

FT Active-site 966 /note= "catalytic triad residue"

XX SG42456-A1.

XX 15-AUG-1997.

XX 27-JUL-1995; 95SG-00000963.

XX 19-AUG-1994; 94US-00296014.

XX (UYSI-) UNIV SINGAPORE NAT.

XX Ling DJ, Bow H;

XX WPI; 1997-557571/51.

DR N-PSDB; AAV06055.

XX Recombinant Carinoscorpis rotundicauda factor C proteins - useful in
PT bacterial endotoxin assays.
XX Claim 3; Page 67-75; 124pp; English.
CC This protein comprises a factor C proenzyme of Singapore horseshoe crab
CC (Carinoscorpis rotundicauda). Its amino acid sequence was deduced from
CC cDNA clone CrFC 21 (see AAV06055), isolated from an amoebocyte cDNA
CC library. A variant form (see AAW43393) of factor C is encoded another
CC newly isolated cDNA clone. CrFC 21 (see AAV06054). Also claimed are
CC expression vectors containing DNA encoding factor C, transformed host
CC cells, methods of producing and purifying factor C zymogens, and methods
CC of protecting factor C zymogens from autoactivation by Gram-negative
CC bacterial endotoxin while the proenzyme is being purified and/or
CC processed from amoebocyte lysates or from recombinant clones. Factor C is
CC useful in endotoxin assays

```
XX
SQ      Sequence 1019 AA;
      Query Match      93.0%; Score 53; DB 2; Length 1019;
      Best Local Similarity 75.0%; Pred. No. 11;
      Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 DXCXGDSGGPXV 12
      | | | | | | | |
Db      960 DACSGDSGGPLV 971
```

Search completed: November 1, 2005, 23:42:20
Job time : 63.2353 secs